

The topology of the IRAS Point Source Catalogue Redshift survey

Article (Published Version)

Canavezes, A, Springel, V, Oliver, S J, Rowan-Robinson, M, Keeble, O, White, S D M, Saunders, W, Efstathiou, G, Frenk, C S, McMahon, R G, Maddox, S, Sutherland, W and Tadros, H (1998) The topology of the IRAS Point Source Catalogue Redshift survey. *Monthly Notices of the Royal Astronomical Society*, 297 (3). pp. 777-793. ISSN 0035-8711

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/89695/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

The topology of the *IRAS* Point Source Catalogue Redshift Survey

A. Canavezes,¹ V. Springel,² S. J. Oliver,¹ M. Rowan-Robinson,¹ O. Keeble,¹
S. D. M. White,² W. Saunders,³ G. Efstathiou,⁴ C. S. Frenk,⁵ R. G. McMahon,⁴
S. Maddox,⁴ W. Sutherland⁶ and H. Tadros⁷

¹Imperial College of Science, Technology and Medicine, Blackett Laboratory, Prince Consort Road, London SW7 2BZ

²Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85740 Garching bei München, Germany

³Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ

⁴Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA

⁵Department of Physics, University of Durham, South Road, Durham DH1 3LE

⁶Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH

⁷Department of Physics, University of Sussex, Falmer, Brighton BN1 9QH

Accepted 1998 February 5. Received 1998 February 5; in original form 1997 July 11

ABSTRACT

We investigate the topology of the new Point Source Catalogue Redshift Survey (PSCz) of *IRAS* galaxies by means of the genus statistic. The survey maps the local Universe with approximately 15 000 galaxies over 84.1 per cent of the sky, and provides an unprecedented number of resolution elements for the topological analysis. For comparison with the PSCz data we also examine the genus of large N -body simulations of four variants of the cold dark matter (CDM) cosmogony. The simulations are part of the Virgo project to simulate the formation of structure in the Universe. We assume that the statistical properties of the galaxy distribution can be identified with those of the dark matter particles in the simulations. We extend the standard genus analysis by examining the influence of sampling noise on the genus curve and introducing a statistic able to quantify the amount of phase correlation present in the density field, the *amplitude drop* of the genus compared to a Gaussian field with identical power spectrum. The results for PSCz are consistent with the hypothesis of random-phase initial conditions. In particular, no strong phase correlation is detected on scales ranging from 10 to 32 h^{-1} Mpc, whereas there is a positive detection of phase correlation at smaller scales. Among the simulations, phase correlations are detected in all models at small scales, albeit with different strengths. When scaled to a common normalization, the amplitude drop depends primarily on the shape of the power spectrum. We find that the constant-bias standard CDM model can be ruled out at high significance, because the shape of its power spectrum is not consistent with PSCz. The other CDM models with more large-scale power all fit the PSCz data almost equally well, with a slight preference for a high-density τ CDM model.

Key words: surveys – galaxies: clusters: general – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

All-sky redshift surveys of galaxies selected from the *IRAS* catalogues have had a dramatic impact on our understanding of the large-scale structure and dynamics of the local Universe. The QDOT survey (Lawrence et al., in preparation) provided early evidence that the large-scale clustering

of galaxies was incompatible with the standard cold dark matter (CDM) model (Efstathiou et al. 1990; Saunders et al. 1991). Comparison of our Local Group motion, or the local velocity field as determined from peculiar velocity studies, with the velocity field inferred from the *IRAS* galaxy density has provided determinations of $\Omega^{0.6}/b$ on very large scales (Rowan-Robinson et al. 1990; Kaiser et al. 1991;

Strauss et al. 1992b; Dekel et al. 1993; Nusser & Davis 1994). These dynamical studies were in general compatible with $\Omega=1$ cosmologies with only small deviations of the bias parameter from $b=1$. This suggested that *IRAS* galaxies are reasonably faithful tracers of the mass field, as compared to, e.g. optically selected galaxies. The main advantage of *IRAS* galaxy redshift surveys is, however, the uniform all-sky coverage as a result of the *IRAS* survey planning and the reduced Galactic obscuration in the far-infrared. *IRAS* galaxy surveys were also able to probe cosmologically significant volumes efficiently due to their broad selection function.

The two most significant *IRAS* redshift surveys to date employed two different strategies. The QDOT survey consists of 2387 galaxies randomly selected at a rate of 1 in 6 to the full depth of the *IRAS* Point Source Catalog (PSC), 0.6 Jy, thus exploring the maximum volume at the expense of sampling density. The 1.2-Jy survey (Fisher et al. 1995), an extension of the 2-Jy survey (Strauss et al. 1992a), comprises 5321 galaxies fully sampled from the PSC, but to a shallower depth.

The sparse sampling strategy of QDOT was ideal for low-order statistics such as the galaxy power spectrum, whereas the 1.2-Jy survey was more useful in measurements of higher order statistics which require higher sampling densities (Bouchet et al. 1993; Fisher et al. 1994; Nusser, Dekel & Yahil 1995). Some important studies such as the inferred local velocity field, the convergence of the dipole and the topology greatly benefit from both high sampling densities and cosmologically significant volumes. For these reasons we have constructed the PSCz (Saunders et al. 1994), a complete redshift catalogue of 15 000 galaxies to the full depth of the PSC (0.6 Jy).

The construction of the parent sample for the PSCz was similar to that of the QMW *IRAS* Galaxy Catalogue (Rowan-Robinson et al. 1990) with substantial efforts made to improve the sky coverage, uniformity and completeness. Details are provided in Saunders et al. (in preparation). Optical counterparts were identified using digitized plate material from APM and COSMOS. Galactic sources were excluded on the basis of *IRAS* colours and literature searches. Existing surveys and literature material provided around two-thirds of the redshifts, allowing the full redshift catalogue to be constructed with realistic amounts of telescope time. We have now measured around 5000 new redshifts to provide a final catalogue with redshift completeness of 98 per cent; the observing programme and data reduction are described in Keeble et al. (in preparation).

There has been extensive discussion in the literature on the ability of topological studies to discriminate between different models of structure formation – not only between inflationary models (with random-phase initial conditions) and non-Gaussian models of structure formation, like cosmic strings or textures, but also between models with different dark matter content (Gott & Park 1996; Matsubara & Suto 1996; Kerscher et al. 1997). From these previous studies, it is clear that the genus statistic is a robust discriminator as long as low-noise data are used. Our use of very large simulations of CDM variants in conjunction with the PSCz survey is ideal to reveal the power of this statistic in

discriminating between different models of structure formation.

Topological studies have also been aimed to drawing conclusions on the shape of the power spectrum. Moore et al. (1992) concluded that the amplitude of the genus curves on large scales was inconsistent with the predictions of a constant bias standard CDM model and that the power spectrum was best fitted by a power law with index $n = -1$. Vogeley et al. (1994) showed that the values for the amplitude of the genus curves were again inconsistent with a constant-bias standard CDM model but consistent with an open CDM model, and Protogeros & Weinberg (1997) concluded that the topology of the 1.2-Jy redshift survey was best fitted by a power spectrum with index $n = -1$. However, the quantitative significance of these results was not up to what one would wish, given the limitations of the data they used.

This paper forms part of a sequence describing the first results of the PSCz survey: Saunders et al. (in preparation) describe the cosmography and analyse the counts-in-cells statistics; Rowan-Robinson et al. (in preparation) examine the convergence of the density and velocity dipole; Sutherland et al. (in preparation) discuss the power spectrum; Tadros et al. (in preparation) perform a spherical harmonic analysis of redshift space distortions, while Keeble et al. (in preparation) compare the radial and transverse components of the correlation function. Here we analyse the topology of the PSCz density field: Section 2 discusses the topological statistics applied, while Section 3 describes the construction of the PSCz density map. We present results for PSCz in Section 4, and give an account of the theoretical models used to compare with the data in Section 5. Finally, we present our conclusions in Section 6.

2 TOPOLOGICAL METHODS

2.1 Genus statistics

The statistical tools currently used to confront theories of structure formation with observational data go well beyond the simple two-point correlation function. The search for such ‘higher order’ information is based on the realization that the two-point correlation function (or equivalently the power spectrum) exhausts the statistical content of a system only when this has a Gaussian nature. It is clear, however, that non-linear gravitational clustering will inevitably introduce non-Gaussian features, even allowing for random-phase initial conditions.

In order to measure departures from *Gaussianity*, Gott, Melott & Dickinson (1986) suggested studying the topology of isodensity surfaces and quantifying it via the *genus*. Subsequently, the genus statistic has been applied to a number of redshift surveys (Gott et al. 1989; Moore et al. 1992; Park, Gott & da Costa 1992; Rhoads, Gott & Postman 1994; Vogeley et al. 1994; Protogeros & Weinberg 1997), and has been subject to several theoretical investigations (Hamilton, Gott & Weinberg 1986; Weinberg, Gott & Melott 1987; Melott, Weinberg & Gott 1988; Park & Gott 1991; Beaky, Scherrer & Villumsen 1992; Matsubara & Yokoyama 1994, 1996; Gott, Cen & Ostriker 1996; Matsubara 1996; Matsubara & Suto 1996). The genus has also been considered in

the more general framework of Minkowski functionals (Mecke, Buchert & Wagner 1994).

Given an isodensity surface S , the genus G of that surface can be defined as

$$G(S) = \# \text{ of holes} - \# \text{ of isolated regions} + 1, \quad (1)$$

i.e., a spherical surface has genus 0; a torus has genus 1, whereas a distribution of N disjoint spherical surfaces gives rise to $G = -(N - 1)$.

A more formal definition of the genus can be given by means of the Gauss–Bonnet theorem, which relates the curvature of the surface to the genus. Let $k = (a_1 a_2)^{-1}$ be the local Gaussian curvature of a two-dimensional surface S , i.e., a_1 and a_2 are the two principal radii of curvature. Then the Gauss–Bonnet theorem states that

$$G(S) = -\frac{1}{4\pi} \int_S k \, dA + 1. \quad (2)$$

Defining S as the boundary surface between two regions above and below a threshold density ρ_c , we can calculate $G(S)$ and plot it as a function of that threshold, thus obtaining the *genus curve* of a given density field.

What do we expect the genus curve to look like? If a high threshold is selected, only a few very dense and isolated regions will be above this density value, and the genus is negative. If a low threshold density is chosen, only a few isolated voids are identified and, again, the genus is negative. On the other hand, for a threshold around the mean density value, one expects, in general, that the isodensity surfaces will have a multiply connected structure that resembles a sponge, with a resulting positive genus.

The simplest case occurs, of course, for a Gaussian random field. In this case, the statistical properties of the density field are completely described by the power spectrum, and so the genus will not provide any additional information. Moreover, underdense or overdense regions are statistically indistinguishable. This means the genus curve is symmetric about the mean, which is characteristic of the so-called ‘sponge-like’ topology. Parametrizing the threshold density by the number $v \equiv (\rho - \rho_{\text{mean}})/\sigma_\rho$ of standard deviations from the mean density, and introducing the genus per unit volume $g(S) \equiv [G(S) - 1]/V$, one finds for a Gaussian random field (Hamilton, Gott & Weinberg 1986):

$$g(v) = N(1 - v^2) e^{-v^2/2}. \quad (3)$$

Here the amplitude

$$N = \frac{1}{(2\pi)^2} \left(\frac{\langle k^2 \rangle}{3} \right)^{3/2} \quad (4)$$

is a constant which depends via its second moment

$$\langle k^2 \rangle = \frac{\int k^2 \hat{P}(k) \, d^3k}{\int \hat{P}(k) \, d^3k} \quad (5)$$

on the power spectrum $\hat{P}(k)$ of the (smoothed) density field.

Hence the genus curve of a Gaussian random field exhibits a universal w-shape. Only the genus amplitude depends on the shape of the power spectrum via equation (5), but it is independent of the normalization of $\hat{P}(k)$.

For a non-Gaussian density field, different topologies are expected. For example, the genus curve can be shifted towards a ‘meatball’ topology or towards a ‘Swiss-cheese’ topology. In a meatball topology, the genus curve peaks at a negative value of v , because isolated structures dominate over an extended range at positive densities, whereas in a Swiss-cheese topology, the genus curve peaks at a positive value of v , since the topology is dominated by empty voids up to a larger density value.

The genus curve defined as a function of v has the remarkable property that it is invariant during the linear growth of density fluctuations in the Universe. The density contrast $\delta\rho/\rho$ grows and contours of ρ change, but contours of v are fixed. This is a very important property, as by measuring the genus curve on scales where non-linear growth has yet to occur we can recover information about the primordial density field and, eventually, be able to distinguish between Gaussian and non-Gaussian initial conditions.

As an alternative to labelling the isodensity surfaces with the number of standard deviations from the mean, it is common to parametrize them with the fraction f_{vol} of the survey volume above the given density threshold or, equivalently, by the number v_f given implicitly by

$$f_{\text{vol}} = \frac{1}{(2\pi)^{1/2}} \int_{v_f}^{\infty} e^{-t^2/2} \, dt. \quad (6)$$

Note that the two parametrizations coincide for the case of a Gaussian random field. The volume fraction parametrization has the disadvantage that it is less sensitive to skewness in the density probability distribution. In addition, it does not discriminate between Gaussian distributions and any other distribution which is a one-to-one transformation of the former, for example, a lognormal distribution (Coles & Jones 1991). This is because the contours of constant volume fraction are invariant under such mappings, as well as the topology of the isodensity contours themselves.

As Vogeley et al. (1994) argue, this invariance property is indeed desired if one is not interested in the positive skewness developing during gravitational collapse, but rather in the topological properties of the initial density field. By using the volume fraction to label isodensity contours, the genus curve becomes independent of the one-point probability distribution function (PDF). In this way the topological analysis is not mixed up with the one-point PDF, which can be better studied by other means. Subscribing to this philosophy, we will hereafter use the volume fraction parametrization.

2.2 Calculation of the genus curve

We want to utilize the genus to study the topology of the observed galaxy distribution, which comes in the form of a point set. As a first step, we therefore need to adopt a method to compute suitable surfaces from the point distribution. Mecke et al. (1994) assigned to each galaxy site a ball of radius r , and examined the genus of the union set of these spheres. We will employ the more widely used approach of constructing smoothed maps of the PSCz density field (see Section 3) and considering isodensity contours for the genus statistic.

Our code to calculate the genus curve of a given smoothed density field is based on the algorithm first proposed by Gott et al. (1986) and on the `CONTOUR` code by Weinberg (1988). We have written a new implementation in C, able to compute high-resolution genus curves very quickly. Here a major increase in speed was achieved by sorting the discrete density field first in order to be able to find the threshold values ρ_t for the required values of f_{vol} at no computational expense.

Given a particular surface of constant density, we approximate it by a network of polygonal faces. This procedure does not change the global topology of the surface as long as the grid size is much smaller than the smoothing length.

When such a polygon network is used to approximate a compact surface of genus G , it can be shown (Gott et al. 1986) that the genus is given by

$$\sum_i D_i = 4\pi(1 - G), \quad (7)$$

where $D_i = 360^\circ - \sum_j V_j$ is the angle deficit at each vertex, and V_j are the angles around the vertex. For example, in a cube, which approximates a compact surface of genus 0, there are three squares around each vertex, and thus $D_i = 360^\circ - 3 \times 90^\circ = 90^\circ$ at each of the eight vertices, giving $\sum_i D_i = 720^\circ = 4\pi$, as expected. The reason for this is that the curvature is compressed into δ -functions at the vertices. Parallel transport arguments show that the integral of the δ -functions of Gaussian curvature over the infinitesimal area of a vertex is just equal to the angle deficit at that vertex.

In practice, we divide space into cubic lattices for sampling either galaxy or simulation data; we then calculate the genus of a given isodensity surface by approximating it by a network of square faces, and by adding up the angle deficits of all vertices.

2.3 Genus-related statistics

In order to measure departures of the observed genus curve from the random-phase shape, we use the genus meta-statistics introduced by Vogeley et al. (1994). They consist of an appropriately defined amplitude, width, and shift of the genus curve. Additionally, we consider the *amplitude drop* of the genus curve compared to the random-phase expectation.

2.3.1 Amplitude

We measure the amplitude of the genus curve as the amplitude N of the best-fitting random-phase genus curve by minimizing χ^2 in the range $-1 < v < 1$. If the underlying density field is sufficiently close to Gaussian, the amplitude N provides direct information on the shape of the power spectrum. It should be noted, however, that the amplitude is systematically biased high by shot noise due to finite sampling of the density field, as will be discussed below.

2.3.2 Width

The second of the meta-statistics is the width of the genus curve, which refers to the range of v over which the genus is

positive, i.e.,

$$W_v = v_+ - v_-, \quad (8)$$

where v_+ and v_- are the first zero crossings of the genus curve, right and left of the origin. The width W_v measures whether the examined density field is more or less sponge-like than a random-phase field, which has $W_v = 2$. In order to improve the reliability of the determination of v_+ and v_- , we first reduce the noise in the genus curve by boxcar smoothing it with a filter of total width $\Delta v = 0.2$. This smoothing is applied only for this statistic.

2.3.3 Shift

In order to distinguish between a Swiss-cheese and a meatball topology, we use a measure for the shift of the peak of the genus curve. this shift may be quantified via

$$\Delta v = \frac{\int v g(v)_{\text{obs}} dv}{\int g(v)_{\text{fit}} dv}, \quad (9)$$

where $g(v)_{\text{obs}}$ is the observed genus curve, and $g(v)_{\text{fit}}$ is the corresponding best-fitting random-phase curve (Park et al. 1992). We evaluate Δv in the range $-1 < v < 1$. A value of $\Delta v > 0$ indicates a bubble-shift, whereas $\Delta v < 0$ means a meatball-shift.

2.3.4 Amplitude drop

With the above meta-statistics, any positive detection of a deviation of the genus curve from the random-phase shape signals the presence of non-Gaussian features in the density field. However, the converse is not necessarily true. For example, we find for the N -body models analysed below that the genus curve retains its random phase shape almost perfectly at all smoothing scales, even when the smoothed density field exhibits already strong non-Gaussian features that develop in the non-linear growth of structure. However, the amplitude of these genus curves is reduced compared to the expectation based on the power spectrum alone.

This amplitude drop can be taken as evidence for phase correlations that develop in the weakly non-linear regime. In order to quantify this effect, we define the amplitude drop

$$R = \frac{N}{N_{\text{rp}}}, \quad (10)$$

where N_{rp} is the amplitude of the random-phase density field with the same power spectrum as the examined density field.

There are different possible ways to measure N_{rp} . For a fully sampled N -body simulation with periodic boundary conditions, perhaps the simplest way is to Fourier-transform the original density field, randomize the phases in Fourier space subject to the condition $\delta_k^* = \delta_{-k}$, thereby *Gaussianizing* the field, and then transform back to real space and measure the genus again. This procedure will be applied to the simulations that we study in Section 5.

For the case of PSCz, there are non-trivial phase correlations between the mask and the density field due to the odd-

shaped survey volume, which spoil this procedure. Instead of constructing a representation of the corresponding random-phase density field, the estimate of N_{rp} needs to be based on a direct measurement of the second moment of the power spectrum.

For this purpose, Vogeley et al. (1994) estimated the power spectrum directly, but this is also complicated by the presence of the mask, and in addition it requires a subtle integration of the measured spectrum to derive N_{rp} .

As an alternative, we propose to measure the variance $\sigma_i^2(\lambda)$ of the underlying density field as a function of smoothing length. This allows an estimate of $\langle k^2 \rangle$ via the logarithmic slope of $\sigma_i^2(\lambda)$:

$$\langle k^2 \rangle \lambda^2 = - \frac{d \log \sigma_i^2}{d \log \lambda}. \quad (11)$$

From this the amplitude N_{rp} can be obtained using equation (4). Our estimator for the variance $\sigma_i^2(\lambda)$ is discussed in Appendix A. In order to reduce the cosmic variance in the estimate of $\langle k^2 \rangle \lambda^2$, we actually fit the measured variances with an analytic form based on a generic CDM power spectrum, and compute the derivative from this fit.

2.4 Error estimates

2.4.1 Uncertainty due to finite sampling

We assume that the observed galaxy distribution represents a Poisson realization of the underlying density field. Since we have only one sampling realization of the PSCz density field at our disposal, we employ the bootstrap technique to derive estimates of the uncertainties due to sampling noise in the PSCz genus results.

For this purpose we generate an ensemble of bootstrap resamplings of the PSCz galaxy catalogue by randomly redistributing the galaxies on to the sites provided by the original catalogue. Note that certain positions will be left empty, while others will be occupied by several galaxies. An estimate for the statistical uncertainty of a measured quantity due to finite sampling may then be obtained as the variance among the measurements for the bootstrap ensemble.

In the context of genus statistics this method has been proven to be a reliable and robust – if somewhat conservative – estimator of statistical errors. Moore et al. (1992) compared bootstrap errors for the genus statistic with the statistical errors obtained for different resamplings of full N -body simulations, and found that the bootstrap errors are 10 per cent higher than the ‘true’ statistical uncertainty.

Of course, in addition to the sampling noise the genus curves are also uncertain due to cosmic variance resulting from the finite survey volume accessible to PSCz. We believe that PSCz is the best sample examined so far in this respect, and that due to the large number of resolved structure elements the cosmic variance is smaller than in previous studies. However, perhaps the only reliable way to quantify its influence is to examine PSCz-like mock surveys extracted from large-volume N -body simulations. This technique allows a careful statistical study of systematic effects, and will be applied in a forthcoming paper by Springel et al. (1998).

2.4.2 Shot-noise bias

In addition to the variance introduced by finite sampling, there is also a systematic effect on the genus curve due to shot noise.

In Appendix B we calculate an estimate of this bias. To the extent that the shot noise does not spoil the Gaussian nature of the underlying field, the genus amplitude may be estimated by

$$N = \frac{1}{(2\pi)^2} \left[\frac{\langle k^2 \rangle}{3} + \frac{\sigma_{\text{shot}}^2}{\sigma_i^2 + \sigma_{\text{shot}}^2} \left(\frac{1}{\lambda^2} - \frac{\langle k^2 \rangle}{3} \right) \right]^{3/2}. \quad (12)$$

Here, σ_{shot}^2 is the variance introduced by shot noise on a given scale, and σ_i^2 is the variance in the field free from shot noise on that same scale.

When the power spectrum follows a power law with slope n on the relevant scales, the amplitude becomes

$$N = \frac{1}{(2\pi)^2 \lambda^3} \left[\frac{n+3}{3} + \frac{\sigma_{\text{shot}}^2}{\sigma_i^2 + \sigma_{\text{shot}}^2} \left(\frac{-n}{3} \right) \right]^{3/2}. \quad (13)$$

We tested this relation with Monte Carlo experiments and found it to be fulfilled very well. Since in a realistic application we typically have $-3 < n < 0$, shot noise tends to increase the measured genus amplitude.

In order to quantify the contribution of shot noise to the genus amplitude, we introduce the ratio

$$A_{\text{shot}} = \frac{N_{\text{obs}}}{N_i}, \quad (14)$$

where N_{obs} is the observed genus amplitude under the influence of noise, and N_i is the amplitude of the field free from shot noise.

Since the variance of the density field due to shot noise σ_{shot} varies with distance, we compute an estimate of A_{shot} by averaging relation (12) over the survey volume.

The variances σ_i^2 and $\sigma_{\text{shot}}^2(r)$ are estimated as outlined in Appendix A. We also need the logarithmic slope $\langle k^2 \rangle \lambda^2$ defined in equation (11). In practice, we want to obtain an estimate of A_{shot} that is not strongly affected by fluctuations due to cosmic variance between the different adopted survey volumes. We therefore base the estimate of A_{shot} on values for σ_i^2 derived from an analytic fit to the measured variances and on slopes $\langle k^2 \rangle \lambda^2$ obtained for this fit.

With these quantities computed, the integrations required to estimate A_{shot} can be done. Typically, we find $A_{\text{shot}} \sim 1.1$ – 1.2 for PSCz, so that genus amplitudes are increased by up to 20 per cent. It is therefore crucial to correct amplitudes for shot noise if one is to draw quantitative conclusions from the measured genus amplitude. For this reason, we calculate a corrected amplitude $N^* = N/A_{\text{shot}}$ in order to estimate the amplitude drop R , which is particularly sensitive to the value of the amplitude.

3 CONSTRUCTION OF DENSITY MAPS

3.1 The PSCz redshift survey data

The sky coverage of the PSCz redshift survey is 84.1 per cent, excluding only the zone of avoidance, here defined by

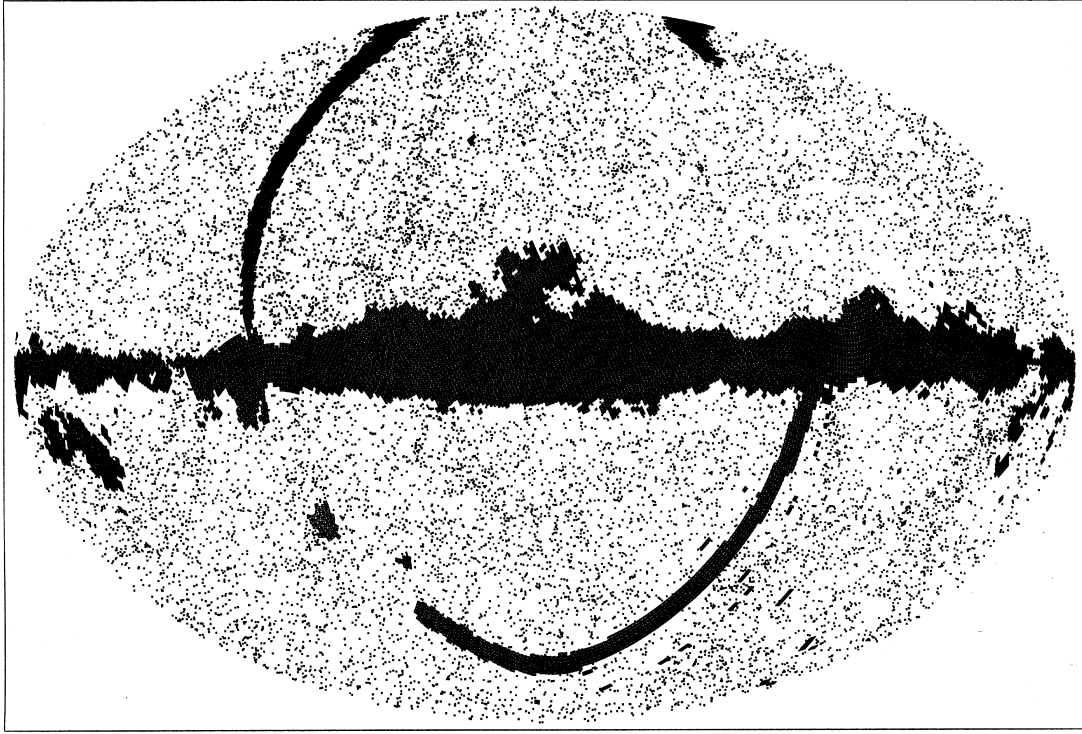


Figure 1. Sky coverage of the PSCz survey in an Aitoff projection. The dots represent the galaxies in the survey; the shaded regions are unobserved and comprise the angular mask. The Galactic Centre lies at the centre of the plot.

an infrared background exceeding 25 MJy sr^{-1} at $100 \mu\text{m}$, and a few unobserved or contaminated patches at higher latitude. The excluded regions are coded in an angular mask, as shown in Fig. 1.

We convert the observed redshifts to the Local Group frame, and use them directly to infer comoving distances without further corrections for peculiar velocities. As has been demonstrated with N -body experiments (Protogeros & Weinberg 1997) and analytical calculations (Matsubara 1996), redshift space distortions hardly affect genus statistics. Therefore this should be a good approximation for the genus analysis. However, a few blueshifted galaxies had to be discarded, resulting in an effective sample of 14 505 galaxies. We will assume an Einstein–de Sitter model for the background cosmology throughout. The results should not be sensitive to this choice because in a cosmological context the PSCz density field maps only the local Universe.

3.2 Smoothing procedure

Assuming a universal luminosity function, an unbiased estimate of the galaxy density field $\rho(r)$ can be obtained by weighting the discrete point distribution $m(r)$ of the observed galaxies with the inverse of the selection function $S(r)$:

$$\rho(r) \propto \frac{m(r)}{S(r)}. \quad (15)$$

Here the selection function $S(z) = \langle m(r) \rangle$ is defined as the mean expected comoving number density of sources at red-

shift z corresponding to the comoving position r . We employ the fitting form

$$S(z) = \frac{\psi}{z^\alpha (1 + (z/z^*)^\gamma)^{\beta/\gamma}}, \quad (16)$$

and determine its parameters (see Table 1) with the methods outlined in Springel & White (1998). Note that the selection function includes a correction for the strong evolution seen in *IRAS* galaxies.

In order to obtain an estimate of the density field smoothed on some scale λ , we convolve $\rho(r)$ with a Gaussian filter of the form

$$W(x) = \frac{1}{\pi^{3/2} \lambda^3} \exp\left(-\frac{x^2}{\lambda^2}\right). \quad (17)$$

Note that here we use Gott et al.'s (1989) definition of a Gaussian filter rather than the conventional form.

However, due to the lack of galaxies in the regions of the angular mask, the density would be systematically underestimated at locations close to unobserved patches if the smoothing were just done by a straightforward use of the

Table 1. Parameters of the selection function of PSCz.

α	β	γ
$0.991^{+0.068}_{-0.073}$	$3.445^{+0.173}_{-0.158}$	$1.925^{+0.162}_{-0.153}$
z^*	$\psi [h^3 \text{Mpc}^{-3}]$	
$0.02534^{+0.00130}_{-0.00116}$	$(141.3 \pm 2.4) \times 10^{-6}$	

kernel of equation (17). In order to avoid this problem, we employ the ratio method proposed by Melott & Dominik (1993), who have shown in a systematic study that a smoothing according to

$$\hat{\rho}(\mathbf{r}) = \frac{\int W(\mathbf{r}-\mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}'}{\int W(\mathbf{r}-\mathbf{r}'') M(\mathbf{r}'') d\mathbf{r}''} \quad (18)$$

leads to the smallest loss or distortion of topological information compared to a number of alternative schemes that treat the mask differently. Here $M(\mathbf{r})$ is a mask field defined to be equal to 0 for \mathbf{r} lying behind the angular mask and to be 1 otherwise. For this choice the denominator of equation (18) essentially renormalizes the smoothing kernel to the survey volume visible from the reference point \mathbf{r} .

In the actual computation of the genus curve we only use the volume with $M(\mathbf{r})=1$ which is not hidden by the mask. Additionally, we restrict the genus computation to a sphere carved out of the smoothed density field. Note that there is no boundary smoothing effect due to the outer surface of this sphere, since we also include the sources outside this final region in the smoothing process.

The method we apply here has the advantage that it does not require some form of *filling* of the unobserved regions. As Melott & Dominik (1993) have shown, simple forms of such fillings that come to mind, like a constant-density padding or randomly placed points of the mean background density, lead to larger biases in the genus curve than the ratio method.

We compute the convolutions that appear in the numerator and denominator of (18) using a Fast Fourier Transform (FFT) on a 128^3 mesh. We choose a grid size of $b=\lambda/6$, which ensures that the genus is free from finite mesh size effects (Hamilton et al. 1986). The final depth R_{\max} of the density field that we use for the topological analysis is always small enough to avoid wrap-around effects due to the periodic FFT smoothing.

3.3 Depth of maps

because the galaxy density of a flux-limited sample declines quickly with distance, the uncertainty of the smoothed density estimate grows rapidly with redshift. It is desirable, of course, to use a survey volume that is as large as possible in order to beat down statistical noise and cosmic variance. However, with respect to the genus statistics the sampling must be at least dense enough to make discreteness effects in the genus curve negligible.

According to a useful rule of thumb (Weinberg et al. 1987), discreteness effects are small if

$$\lambda \geq d = S^{-1/3}, \quad (19)$$

where d is the mean interparticle separation, and S is the selection function. Adopting this criterion, we choose a maximal radius R_{\max} , given by $\lambda = S(R_{\max})^{-1/3}$, and use it to delimit the usable survey volume V_s . This choice ensures that at the far edge of the survey volume the sampling condition is just met, and in the remainder of the volume the sampling is denser.

However, if λ is of the order of d , one risks artificially introducing a meatball bias by identifying lonely tracers as isolated clusters. Any observed bias of this kind should

therefore be treated with caution, since it might well be a discreteness effect.

Additionally, the genus amplitude is biased due to shot-noise present in the reconstructed density field, as we have discussed in Section 2. This noise contributes to second and higher moments of the density field. Unfortunately, the genus depends in a non-linear way on the various moments of the density field, making it extremely difficult to derive an unbiased genus estimator which is not affected by it. Hence we estimate the contribution of shot noise to the genus signal a posteriori.

3.4 Resolution elements

The notion of *number of resolution elements* provides a useful way to roughly compare the statistical power of genus measurements. Because the smoothing extends over an effective volume $V_{\text{sm}} = \pi^{3/2} \lambda^3$, the number of independent structures that can be present in a finite survey volume is limited. This number is of order

$$N_{\text{res}} = \frac{V_s}{V_{\text{sm}}} = \frac{\omega R_{\max}^3}{\pi^{3/2} \lambda^3}, \quad (20)$$

where ω is the solid angle covered by the survey.

The number N_{res} indicates the power of a data set used for topological analysis. With the QDOT survey Moore et al. (1992) reached a maximum of about $N_{\text{res}}=80$, whereas the most powerful genus results so far came from Vogeley et al.'s (1994) treatment of the CfA survey, where they reached $N_{\text{res}}=260$ for their best subsample. In a recent study of the 1.2-Jy redshift survey Protogeros & Weinberg (1997) achieved $N_{\text{res}}=170$.

The PSCz redshift survey can provide still more resolution elements, as is evidenced in Fig. 2 and Table 2. It also provides a high number of resolution elements over an unprecedented wide range of smoothing scales. In particular, there are more than 300 resolution elements in the range $6 \leq \lambda \leq 20 h^{-1}$ Mpc. This wide dynamic range together with the large volume covered make it more powerful than all previously examined samples.

For the genus analysis we have examined smoothing lengths between 5 and $56 h^{-1}$ Mpc with approximately loga-

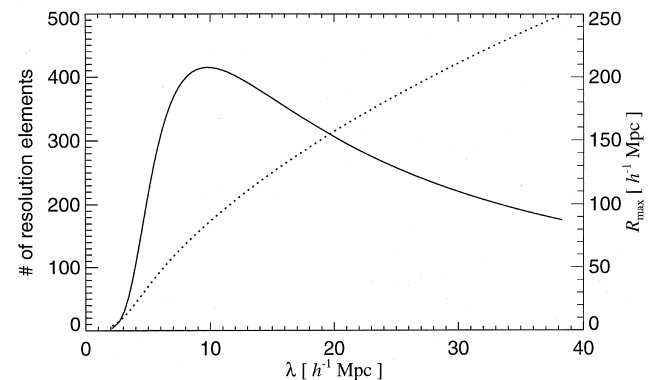


Figure 2. The number of resolution elements (solid, left-hand scale) for the PSCz survey when the maximal survey volume is used. Also shown is the radius (dotted, right-hand scale) of the usable survey volume.

Table 2. The smoothing lengths adopted for the topological analysis of the PSCz survey. Listed are the adopted survey depth R_{max} , the resulting number N_{res} of resolution elements, and the number N_{gal} of galaxies inside the survey volume.

λ [$h^{-1}\text{Mpc}$]	R_{max} [$h^{-1}\text{Mpc}$]	N_{res}	N_{gal}
5	34.92	215.5	2295
6	47.16	307.3	3550
7	58.35	366.3	4928
8	68.58	398.5	5909
10	86.91	415.3	7510
12	103.19	402.3	8775
14	118.04	379.2	9681
16	131.81	353.7	10356
20	156.96	305.8	11309
24	179.79	266.0	11968
28	200.97	233.9	12496
32	221.79	210.6	12833
40	263.44	180.7	13339
48	305.09	162.4	13684
50	315.50	158.9	13748
56	346.73	150.2	13924

rhythmic spacing. In particular, all the smoothing lengths considered in the studies by Moore et al. (1992), Rhoads et al. (1994) and Vogeley et al. (1994) are contained in this set. Table 2 lists some relevant parameters for the different cases.

4 PSCZ RESULTS

4.1 PSCz genus curves

Fig. 3 shows the genus curves of PSCz¹ for smoothing lengths of 6, 8, 12, 16, 24, 32, 48, and 56 $h^{-1}\text{Mpc}$. The other smoothing scales examined show genus curves with a similar appearance. In each panel the solid line shows a high-resolution curve computed for spacing $\Delta v = 0.01$. While it is clear that adjacent points on the genus curve are highly correlated, the amount of jitter present in the curves nicely indicates the level of noise contained in the measurements. Also shown as a dashed line is the best-fitting random-phase genus curve, which we use to infer the measured genus amplitude.

We have chosen not to show the average genus curve over the set of bootstrap realizations (which would be a much smoother curve), because of the high degree of correlation between the different points in the curve. Furthermore, the bootstrap-generated curves are expected to be biased towards a meatball topology. However, we use the bootstrap-generated curves to estimate the statistical uncertainty of individual points on the genus curve. Similarly, we use the bootstrap curves to estimate uncertainties for the genus meta-statistics. For this purpose, we estimate the uncertainty of a particular meta-statistic as the rms fluctuation of the values obtained for the set of 15 bootstrap resamplings of the data. Table 3 summarizes the numerical results obtained for the genus meta-statistics of PSCz.

As is evident from the plots in Fig. 3, the PSCz genus curves follow the random-phase expectation rather well. In particular, no obvious deviations like shifts, broadenings or the like are observed.

This is also confirmed by Figs 4 and 5, which plot the meta-statistics shift and width for scales ranging from 5 to 56 $h^{-1}\text{Mpc}$. The error bars give the rms fluctuations obtained over 15 bootstrap resamplings of the data. A comparison is also given with the results of Vogeley et al. (1994) for the CfA survey.

Apart from a slight indication of a meatball-shift and a broadening of the genus curves at small scales, these two meta-statistics fail to show any significant departures from Gaussianity.

We turn now to the amplitudes of the genus curves themselves. Fig. 6 displays our results for PSCz, together with the amplitudes of variants of CDM models, which we will discuss fully in the next section. Also shown is the expected amplitude in linear theory for the power spectrum (21) with shape parameter $\Gamma = 0.2$. The amplitudes of the simulations are based on the dark matter distribution, which we identify with the galaxy distribution. It should be stressed that the use of volume weighting makes this a relatively weak assumption, since we only need regions of higher smoothed galaxy density to correspond to regions of higher smoothed mass density, but not galaxy and mass densities to be proportional to one another. The use of volume weighting makes the genus curves insensitive to details of bias (Park & Gott 1991).

Based on Fig. 6, the standard cold dark matter (SCDM) model in which the galaxies trace the mass is clearly ruled out at high significance level. This can be traced back to the fact that the shape of its power spectrum is in strong disagreement with PSCz. The other three N -body models do much better in this respect; they match the PSCz amplitudes very well, because the shape of their power spectrum is in good agreement with PSCz.

It is important to point out the significance of these results. Although a counts-in-cell or power spectrum analysis of the PSCz redshift survey will probably lead to the same result for the shape of the power spectrum, the route we followed here is radically different. The genus depends on the spatial coherence of structures rather than on the rms amplitude of fluctuations. Moreover, it depends only on the rank order of density values, not on the values themselves. In the presence of unusual forms of biased galaxy formation, non-Gaussian initial conditions, or observational errors, there is no guarantee a priori that the determination of the shape of $P(k)$ via the genus amplitude will be consistent with that inferred from a direct measurement of $P(k)$. The fact that the same conclusion is reached by a very different analysis of the data lends considerable, largely independent support to the $P(k)$ result.

Note that, in Fig. 6, we have plotted the amplitude N of PSCz without a shot-noise correction, i.e., these values are likely to be biased slightly *high*.

4.2 Variance and amplitude drop

Fig. 7 shows our results for the variance of the Gaussian-smoothed PSCz density fields, obtained with the unbiased estimator described in Appendix A. The error estimates are

¹The PSCz genus curves may be retrieved electronically at <http://www.mpa-garching.mpg.de/PUBLICATIONS/DATA/971121-pscz>

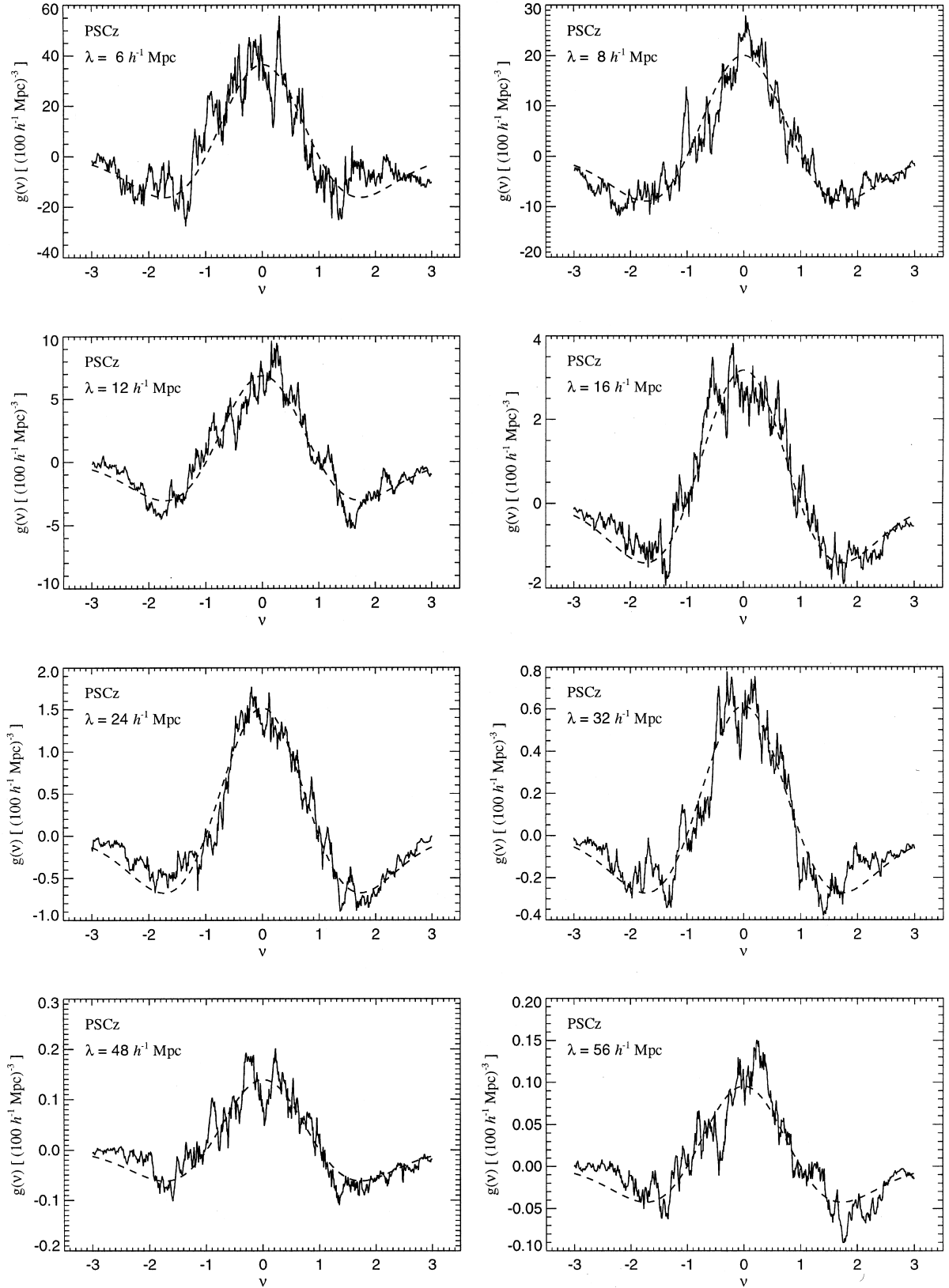


Figure 3. The genus curves of the PSCz redshift survey for selected smoothing lengths. In each panel, the solid line represents the raw PSCz genus, measured at resolution $\Delta v = 0.01$. The dashed line gives the best-fitting Gaussian curve, which also defines the amplitude of the genus curve.

Table 3. Summary of numerical results obtained for the genus meta-statistics of PSCz. Listed for each smoothing length are the genus amplitude in the dimensionless form $(2\pi)\lambda^3 N$, the absolute genus amplitude $G = (4\pi/3)NR_{\text{max}}^3$ in the survey volume, the shift $\Delta\nu$, and the width W_ν . We also give the measured variance σ^2 of the smoothed density fields, our estimated shot-noise correction factors A_{shot} and the amplitude drop R .

λ [h^{-1} Mpc]	$(2\pi)^2\lambda^3 N$	$\Delta\nu$	W_ν	σ^2	A_{shot}	R
5	0.25 ± 0.03	-0.25 ± 0.11	2.25 ± 0.21	0.6691 ± 0.1312	1.11	0.68 ± 0.11
6	0.32 ± 0.03	-0.15 ± 0.09	2.02 ± 0.14	0.6044 ± 0.0772	1.11	0.83 ± 0.11
7	0.34 ± 0.02	-0.07 ± 0.05	2.14 ± 0.13	0.5163 ± 0.0485	1.10	0.83 ± 0.09
8	0.40 ± 0.03	0.05 ± 0.05	2.25 ± 0.15	0.3937 ± 0.0315	1.12	0.90 ± 0.09
10	0.45 ± 0.04	0.01 ± 0.05	2.15 ± 0.17	0.2463 ± 0.0170	1.15	0.91 ± 0.12
12	0.46 ± 0.03	0.03 ± 0.04	2.25 ± 0.10	0.1873 ± 0.0201	1.16	0.85 ± 0.10
14	0.46 ± 0.03	-0.05 ± 0.05	2.04 ± 0.06	0.1474 ± 0.0156	1.16	0.79 ± 0.08
16	0.50 ± 0.05	-0.02 ± 0.07	2.32 ± 0.12	0.1162 ± 0.0122	1.18	0.81 ± 0.12
20	0.63 ± 0.05	0.07 ± 0.06	2.04 ± 0.15	0.0751 ± 0.0079	1.19	0.92 ± 0.11
24	0.82 ± 0.09	0.06 ± 0.05	1.86 ± 0.14	0.0482 ± 0.0055	1.21	1.09 ± 0.19
28	0.87 ± 0.07	0.05 ± 0.05	1.93 ± 0.13	0.0336 ± 0.0042	1.21	1.08 ± 0.16
32	0.83 ± 0.08	0.03 ± 0.07	1.95 ± 0.12	0.0239 ± 0.0033	1.21	0.98 ± 0.20
40	0.75 ± 0.07	-0.05 ± 0.05	1.95 ± 0.19	0.0145 ± 0.0023	1.19	0.82 ± 0.16
48	0.63 ± 0.12	-0.04 ± 0.07	2.08 ± 0.17	0.0103 ± 0.0019	1.16	0.66 ± 0.22
50	0.67 ± 0.12	0.03 ± 0.11	2.06 ± 0.18	0.0100 ± 0.0019	1.15	0.71 ± 0.24
56	0.68 ± 0.15	0.03 ± 0.08	2.15 ± 0.18	0.0089 ± 0.0018	1.12	0.70 ± 0.26

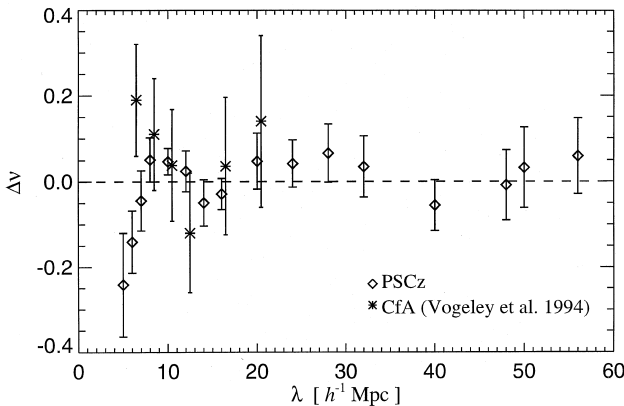


Figure 4. The shift measured for the PSCz genus curves. The error bars are based on bootstrap resamplings of the data. Also shown are the results of Vogeley et al. (1994) for the CfA survey. Here we indicate the uncertainty of CfA by the errors that the authors report for mock catalogues extracted from a LCDM model.

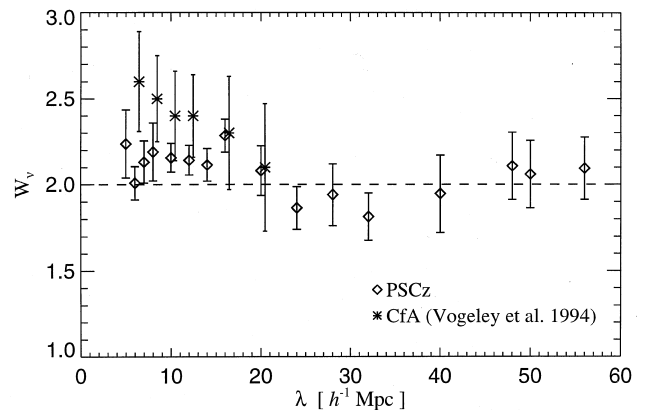


Figure 5. The width of PSCz genus curves, with error estimates derived with the bootstrap technique. Vogeley et al.'s (1994) results for the CfA catalogue are shown for comparison. Here the errors are taken to be the uncertainty that Vogeley et al. (1994) report for mock catalogues extracted from a LCDM model.

based on an ensemble of PSCz-like mock surveys drawn from an N -body simulation. Note that we find somewhat more power on small scales than was found for QDOT (Saunders et al. 1991).

A fit of the linear CDM power spectrum (21) to the measured variances results in a shape parameter $\Gamma = 0.19 \pm 0.04$ and a normalization $\sigma_8 = 0.84 \pm 0.04$, here expressed in terms of the rms fluctuations in top-hat spheres of radius $8 h^{-1}$ Mpc. Note that these values refer to redshift space only; no correction for redshift space distortions has been attempted. The quoted uncertainties are estimated by approximating the measurements for different λ as being independent.

As outlined in Section 2.3.4, we also compute estimates for the logarithmic slope of $\sigma^2(\lambda)$ and the approximate shot-noise correction factor A_{shot} . Combining these measure-

ments allows us to estimate the amplitude drop of PSCz as $R = N^*/N_{\text{rp}} = N/(A_{\text{shot}}N_{\text{rp}})$.

The uncertainty of the measured values for R is difficult to estimate, since we deal with a ratio of correlated quantities. We tentatively assign the relative error of the amplitude measurement as error of R . Our experiments with mock catalogues suggest that this gives approximately the right size of uncertainty.

In Fig. 8 we plot our results for the amplitude drop of PSCz. Although some amount of phase correlation seems to be present at small scales, we find that PSCz does not exhibit strong phase correlations on scales above $10 h^{-1}$ Mpc, in contrast to the findings of Vogeley et al. (1994) for the CfA survey. This means that the PSCz density field is consistent with random-phase initial conditions. In the next section we will further analyse this amplitude drop and compare it to the results of N -body simulations.

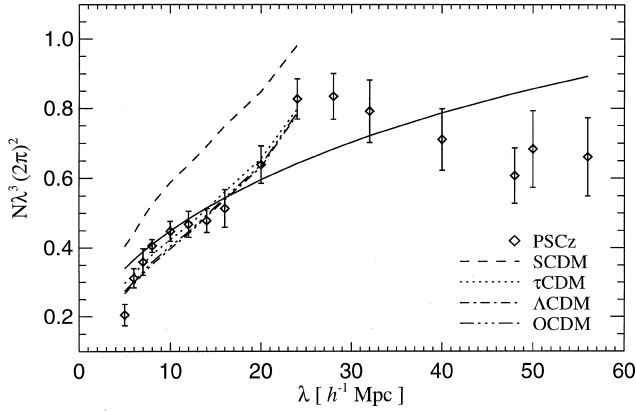


Figure 6. Genus amplitudes for PSCz and for the four CDM models. The error bars for PSCz represent bootstrap estimates of the sampling uncertainty. The solid line gives the expectation based on linear theory for a power spectrum with shape parameter $\Gamma=0.2$. While PSCz matches the τ CDM, Λ CDM and OCDM simulations well, the SCDM model exhibits genus amplitudes that are too large.

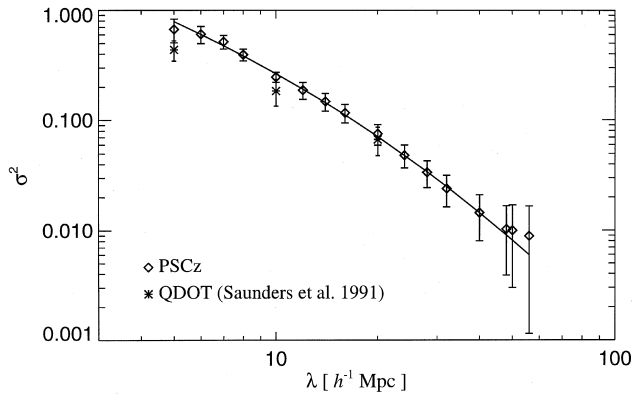


Figure 7. The variance $\sigma^2(\lambda)$ found in Gaussian cells for the PSCz survey. The 1σ uncertainties are estimated from a set of PSCz-like mock catalogues extracted from an N -body simulation. The solid line represents the best fit of the power spectrum (21) to the data, with $\Gamma=0.19 \pm 0.04$ and $\sigma_8=0.84 \pm 0.04$. For comparison, also shown is the result of Saunders et al. (1991) for the QDOT survey (with their error bars).

5 N-BODY SIMULATIONS OF CDM MODELS

5.1 The simulations

We study four variants of the CDM cosmogony. These models were kindly provided by the Virgo collaboration (Colberg et al. 1998; Jenkins et al. 1998), who have recently started an ambitious project to model the formation of structure in the Universe making use of the largest available supercomputers in Europe. The simulations have been performed with an AP³M-SPH code named HYDRA (Couchman, Thomas & Pearce 1995).

All four models we examine contain only CDM in periodic boxes of size $239.5 h^{-1}$ Mpc. The basic parameters of the individual runs are given in Table 4. Here the shape

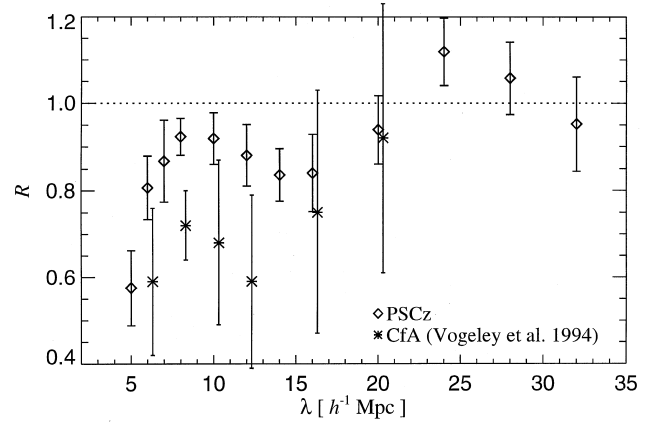


Figure 8. The amplitude drop R measured for PSCz. Also shown are the results obtained by Vogeley et al. (1994) for the CfA survey. Here the errors are taken to be the uncertainty that Vogeley et al. (1994) report for mock catalogues extracted from a Λ CDM model.

Table 4. Parameters of the examined CDM models. The simulations have been done by the Virgo collaboration.

	SCDM	τ CDM	Λ CDM	OCDM
Number of particles	256 ³	256 ³	256 ³	200 ³
Box size [h^{-1} Mpc]	240	240	240	240
z_{start}	50	50	50	50
Ω_0	1.0	1.0	0.3	0.3
Ω_Λ	0.0	0.0	0.7	0.0
Hubble constant h	0.5	0.5	0.7	0.7
Γ	0.5	0.21	0.21	0.21
σ_8	0.60	0.60	0.90	0.85

parameter Γ refers to the linear-theory power spectrum

$$P(k) = \frac{Bk}{\{1 + [ak + (bk)^{3/2} + (ck)^2]^\nu\}^{2/\nu}}, \quad (21)$$

where $a=6.4\Gamma^{-1} h^{-1}$ Mpc, $b=3.0\Gamma^{-1} h^{-1}$ Mpc, $c=1.7\Gamma^{-1} h^{-1}$ Mpc and $\nu=1.13$ (Bond & Efstathiou 1984; Efstathiou, Bond & White 1992). The models are normalized so as to match the observed abundance of rich clusters of galaxies (Eke, Cole & Frenk 1996).

The SCDM model has the power spectrum (21) with a shape parameter $\Gamma=0.5$ in the linear regime, as predicted from a CDM inflationary scenario with a primordial scale-invariant Harrison-Zel'dovich spectrum. The τ CDM model is a variant of CDM with more power on large scales, as suggested by numerous recent observations (Saunders et al. 1991; Peacock & Dodds 1994; Oliver et al. 1996). The same shape of the power spectrum may also be obtained in low-density universes. We consider two models of this class; one open universe (OCDM), and one with a cosmological constant (Λ CDM) that retains a flat background geometry.

The simulations contain typically more than 16.7 million particles, thus providing excellent data which allow very accurate measurements of the genus statistic, especially on small scales where the number of resolution elements is

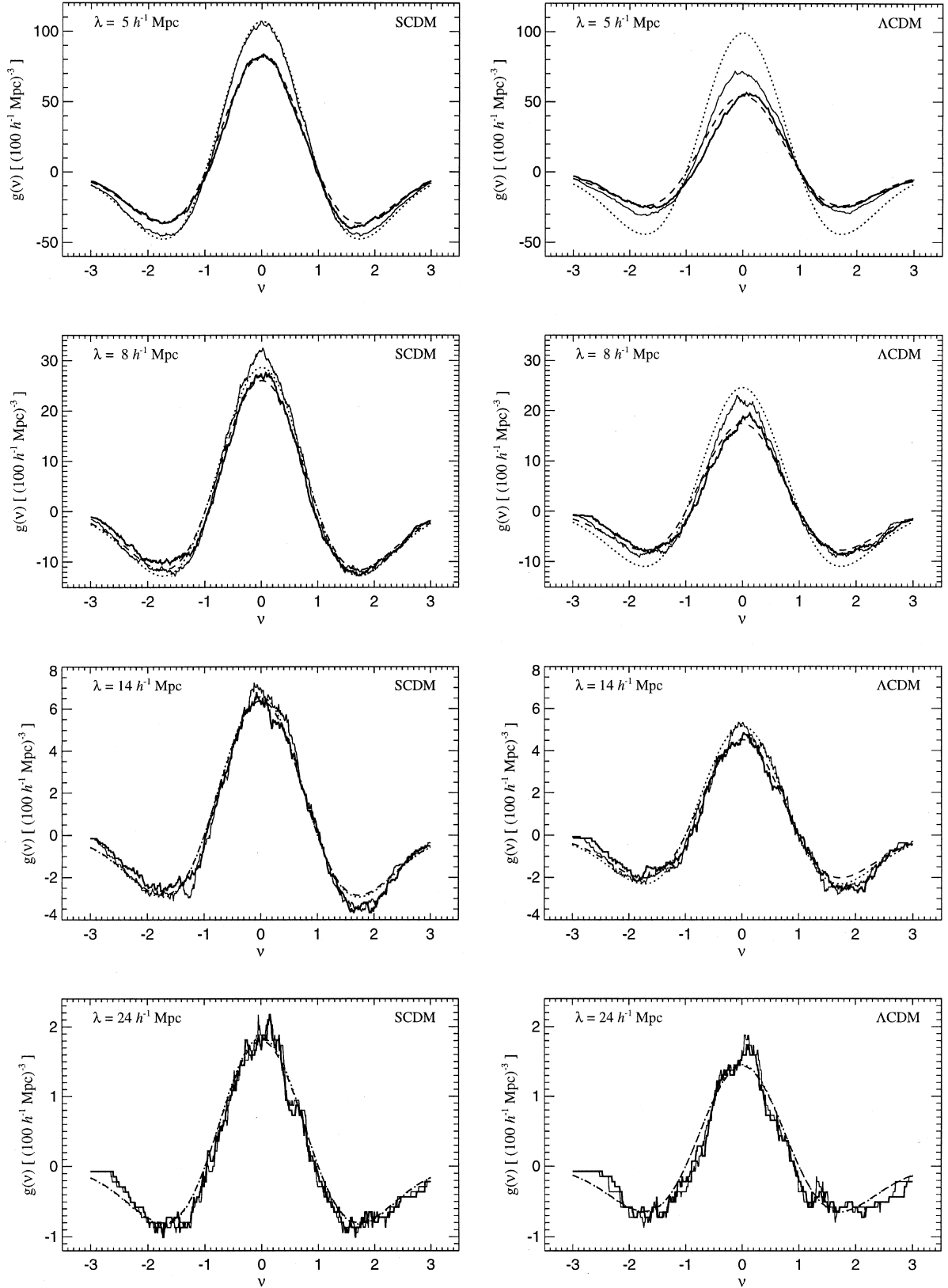


Figure 9. Genus curves of the SCDM (left column) and Λ CDM (right column) simulations at selected smoothing lengths. In each panel we show the actual genus curve of the evolved simulation as a heavy solid line, and we give the best-fitting random-phase genus curve as a dashed line. The dotted line is the genus of the corresponding *Gaussianized* field, while the thin solid line shows the genus curve of the initial conditions.

high and the genus curves are hardly influenced by cosmic variance. This is, however, not true for scales $\geq 20 h^{-1}$ Mpc. The number of resolution elements in the simulations drops from 19 861 for $\lambda = 5 h^{-1}$ Mpc to 180 for $\lambda = 24 h^{-1}$ Mpc. Hence the survey volume is quite limited even for such large simulations. In fact, beyond $20 h^{-1}$ Mpc the volume probed by PSCz becomes larger than that of the simulations. We therefore restrict our genus analysis of the models to the 10 smoothing scales between 5 and $24 h^{-1}$ Mpc used in the analysis of PSCz.

In order to construct smoothed density fields for the simulations, we bin the particles on a 128^3 grid by CIC assignment and subsequently smooth with a Gaussian kernel. We also compute *Gaussianized* representations of the simulations by taking the Fourier transforms of the original density fields, randomizing the phases in Fourier space subject to the reality constraint $\delta_k^* = \delta_{-k}$ and transforming back to real space. These fields are then smoothed in the same way as the original ones, and the genus curves are calculated as before. However, due to the limited simulation volume there is some variance introduced in the resulting genus curves from the variety of different possible random-phase realizations. In order to reduce this effect to a minimum, we compute 10 Gaussianized fields and take the average genus curve of those as the Gaussianized genus curves. The amplitude of the latter is used to estimate the amplitude drop.

5.2 Results

Fig. 9 shows representative genus curves obtained for the SCDM and Λ CDM models. In each panel the heavy solid line gives the genus curve for the density field of the evolved simulation at the indicated smoothing scale. A fit of these data to the generic random-phase genus curve is shown as a dashed line. The dotted line shows the genus curve for the corresponding *Gaussianized* field, and the thin solid line marks the genus of the initial conditions of the simulations. For the sake of brevity, we do not show further genus curves for the τ CDM and OCDM models. They look qualitatively very similar to those in Fig. 9.

Note that there is no sampling noise in these density fields; the residual tremble that is present in the N -body genus curves is caused by the finite number of resolution elements, i.e., ultimately by cosmic variance.

It is apparent that all the genus curves follow the Gaussian w-shape very well. Apart from the small bubble-shift of the Λ -model at the smallest smoothing scales the only meta-statistic able to measure a significant deviation from Gaussianity is the amplitude. It is interesting to note that the Gaussianized genus curves of the SCDM model remain close to the ones of the initial linear density field at all smoothing scales. In contrast, the Λ CDM model shows signs of non-linear evolution of the power spectrum at $\lambda = 5 h^{-1}$ Mpc. Its effect has been to increase the effective slope of the power spectrum (Jenkins et al. 1998) on the relevant scales, and thereby also to increase the genus amplitude of the Gaussianized fields. Higher order correlations, however, ‘keep the genus in place’, and decrease the amplitude at small λ below the linear-theory expectation.

We defined the amplitude drop as the ratio of the genus amplitudes of the evolved density field and its Gaussianized counterpart. Hence this quantity is a quantitative measure

of higher order correlations. In all of the four models an amplitude drop is present at small scales, being largest for the Λ CDM model and smallest for the SCDM model.

In Fig 10 we plot this amplitude drop against the smoothing scale for the four simulations. As expected, the amplitude drop becomes smaller with increasing smoothing length. This is consistent with the picture that all non-linear features are sufficiently smoothed out in this regime. It is interesting, however, that the models exhibit a varying strength of the amplitude drop. For the Λ CDM and OCDM models, phase correlations up to the scale of $24 h^{-1}$ Mpc are detected, whereas the SCDM becomes very close to Gaussian at about $10 h^{-1}$ Mpc, and the τ CDM takes an intermediate position.

Fig. 11 shows that part of this difference is caused by the different normalizations of the models. Here we have plotted the amplitude drop against the variance $\sigma^2(\lambda)$ of the density fields. The good match between the models with shape parameter $\Gamma = 0.21$ suggests that the strength of the

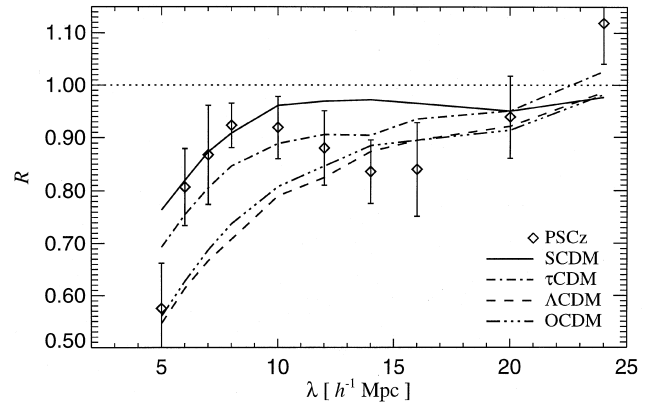


Figure 10. The amplitude drop measured for the four N -body simulations as a function of the smoothing scale. Also shown are the values estimated for the PSCz survey (and error bars).

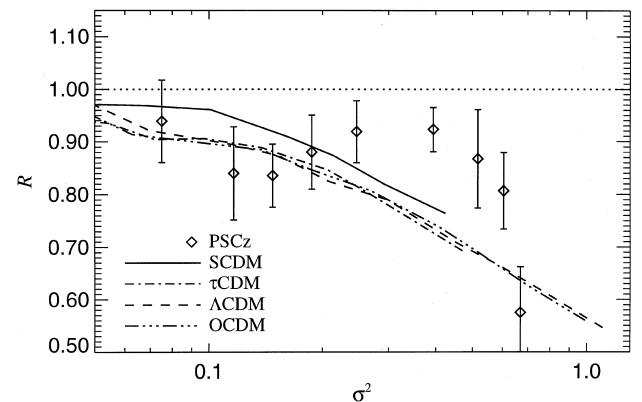


Figure 11. The amplitude drop revisited. Here we plot the data of Fig. 10 again, but this time as a function of the variance of the smoothed density fields. The good match of the three $\Gamma = 0.21$ models in this representation suggests that the differences that show up in Fig. 10 between these models are only due to their slightly different normalizations. The SCDM model ($\Gamma = 0.5$), however, continues to show a considerably smaller amplitude drop.

amplitude drop at fixed σ is primarily governed by the shape of the power spectrum.

In Figs 10 and 11 we have also plotted the measured amplitude drops for the PSCz survey. It is encouraging that the measurements for PSCz exhibit about the right strength of phase correlations like the one expected in CDM scenarios. In principle, one could have hoped that this measurement might be able to favour strongly one of the CDM variants. However, at this point the discriminative power of the genus test is not strong enough to do this. The Λ CDM, τ CDM and OCDM fit the PSCz almost equally well. Only Fig. 10 suggests a slight preference for τ CDM, essentially favouring its normalization. Given the considerable uncertainties in the biasing issue, this is not more than circumstantial evidence for this model. Note that, in contrast to Fig. 10, the horizontal positions of the PSCz points in Fig. 11 depend on bias; a positive bias factor would reduce the mismatch between PSCz and the models at large σ^2 .

It should be noted, however, that we are comparing observations with *perfect* simulation data. Mock catalogues should be used to assess systematic effects accurately from shot noise, redshift-space distortions, selection function, and so forth, so that one can establish firm conclusions about the viability of models. Protogeros & Weinberg (1997) have pioneered this approach, and it will also be applied in a forthcoming paper by Springel et al. (1998).

6 CONCLUSIONS

We have analysed the topology of a new large redshift survey, the PSCz, by means of genus statistics. The genus test provides not only information about the shape of the power spectrum, but is also sensitive to higher order correlations of the density field. The PSCz survey is well suited for the genus test, due to its large survey volume and its near full sky coverage. In particular, it allows a topological analysis with larger statistical significance than any previously examined sample.

The genus curves of PSCz are featureless at smoothing scales ranging from 5 to 56 h^{-1} Mpc, i.e., they exhibit the w-shape that is characteristic of random-phase density fields. In particular, we find no clear evidence for shifts or broadenings of the genus curves. Evidence for significant non-Gaussian signatures on large scales found previously in Vogeley et al. (1994) are not confirmed by PSCz.

When the genus amplitude is examined, we find that the PSCz density field is consistent with a CDM power spectrum with shape parameter $\Gamma \approx 0.2$, and inconsistent with the SCDM model. These results seem consistent with previous analysis of the shape of the power spectrum based on topology (Moore et al. 1992; Vogeley et al. 1994; Protogeros & Weinberg 1997). We expect that an ordinary power spectrum analysis of PSCz will give a similar result. Again we want to stress that these two methods of estimating the shape of the power spectrum are largely independent.

We have demonstrated that the genus curves of CDM models retain the w-shape of the random-phase genus curve even at small smoothing lengths where non-linear gravitational clustering has already generated significant skewness of the one-point PDF. However, the non-linear evolution manifests itself in a depressed genus amplitude compared to

the expectation based on the power spectrum alone. This amplitude drop directly quantifies the amount of higher order correlations present in the density field. When plotted against the variance of the smoothed density field, we find that the amplitude drop depends on the shape of the power spectrum alone.

The measured amplitude drop for PSCz is consistent with the CDM models, i.e., the non-detection of strong phase correlations in PSCz for smoothing scales above 10 h^{-1} Mpc supports the hypothesis that structure grew from random-phase initial conditions. Alternative models for structure formation that provide a non-Gaussian seed field (e.g., by means of cosmic strings or explosions) are expected to exhibit stronger higher order correlations in their density fields. Future analysis of the genus of such models should be used to test whether these correlations indeed show up as a strong amplitude drop or a distortion of the shape of the genus curve. In this case, our results for PSCz can place strong limits on the viability of such models.

ACKNOWLEDGMENTS

We thank the referee, David Weinberg, for many helpful comments on this paper. AC acknowledges the support of JNICT (Portugal). CSF acknowledges a PPARC Senior Research Fellowship. We are grateful to the Virgo consortium (J. Colberg, H. Couchman, G. Efstathiou, C. S. Frenk, A. Jenkins, A. Nelson, J. Peacock, F. Pearce, P. Thomas and S. D. M. White) for providing their N -body simulations in advance of publication, and to Jörg Colberg for many interesting discussions.

REFERENCES

- Beaky M. M., Scherrer R. J., Villumsen J. V., 1992, *ApJ*, 387, 443
- Bond J. R., Efstathiou G., 1984, *ApJ*, 285, L45
- Bouchet F. R., Strauss M. A., Davis M., Fisher K. B., Yahil A., Huchra J. P., 1993, *ApJ*, 417, 36
- Colberg J. M. et al., 1998, in Hamilton D., ed., *Ringberg Workshop, Large-scale Structure in the Universe*. Kluwer, Dordrecht, in press
- Coles P., Jones B., 1991, *MNRAS*, 248, 1
- Colley W. N., Gott R. J., Park C., 1996, *MNRAS*, 281, L82
- Couchman H. M. P., Thomas P. A., Pearce F. R., 1995, *ApJ*, 452, 797
- Dekel A., Bertschinger E., Yahil A., Strauss M. A., Davis M., Huchra J. P., 1993, *ApJ*, 412, 1
- Efstathiou G., Kaiser N., Saunders W., Lawrence A., Rowan-Robinson M., Ellis R. S., Frenk C. S., 1990, *MNRAS*, 247, 10p
- Efstathiou G., Bond J. R., White S. D. M., 1992, *MNRAS*, 258, 1p
- Eke V. R., Cole S., Frenk C. S., 1996, *MNRAS*, 282, 263
- Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23
- Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J., 1994, *MNRAS*, 266, 50
- Fisher K. B., Huchra J. P., Strauss M. A., Davis M., Yahil A., Schlegel D., 1995, *ApJS*, 100, 69
- Gott J. R., Melott A. L., Dickinson M., 1986, *ApJ*, 306, 341
- Gott J. R. et al., 1989, *ApJ*, 340, 625
- Gott J. R., Cen R., Ostriker J. P., 1996, *ApJ*, 465, 499
- Hamilton A. J. S., Gott J. R., Weinberg D., 1986, *ApJ*, 309, 1
- Jenkins A. R. et al., 1998, *ApJ*, in press

- Kaiser N., Efstathiou G., Saunders W., Ellis R., Frenk C., Lawrence A., Rowan-Robinson M., 1991, MNRAS, 252, 1
- Kerscher M. et al., 1997, MNRAS, 284, 73
- Matsubara T., 1996, ApJ, 463, 409
- Matsubara T., Suto Y., 1996, ApJ, 460, 51
- Matsubara T., Yokoyama J., 1994, ApJ, 434, L43
- Matsubara T., Yokoyama J., 1996, ApJ, 457, 13
- Mecke K. R., Buchert T., Wagner H., 1994, A&A, 288, 697
- Melott A. L., Dominik K. G., 1993, ApJS, 86, 1
- Melott A. L., Weinberg D. H., Gott J. R., 1988, ApJ, 328, 50
- Moore B. et al., 1992, MNRAS, 256, 477
- Nusser A., Davis M., 1994, ApJ, 421, L1
- Nusser A., Dekel A., Yahil A., 1995, ApJ, 449, 439
- Oliver S. J. et al., 1996, MNRAS, 280, 673
- Park C., Gott J. R., 1991, ApJ, 378, 457
- Park C., Gott J. R., Melott A. L., Karachentsev I. D., 1992, ApJ, 387, 1
- Park C., Gott J. R., da Costa L. N., 1992, ApJ, 392, L51
- Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020
- Protogerios Z. A. M., Weinberg D. H., 1997, ApJ, 489, 457
- Rhoads J. E., Gott J. R., Postman M., 1994, ApJ, 421, 1
- Rowan-Robinson et al., 1990, MNRAS, 247, 1
- Saunders W. et al., 1991, Nat, 349, 32
- Saunders W. et al., 1994, in Maddox S. J., Aragón-Salamanca A., eds, 35th Herstmonceux Conference, Wide Field Spectroscopy and the Distant Universe. World Scientific, Singapore, p. 88
- Springel V., White S. D. M., 1998, MNRAS, in press
- Springel V. et al., 1998, MNRAS, in press
- Strauss M. A., Huchra J. P., Davis M., Yahil A., Fisher K. B., Tonry J., 1992a, ApJS, 83, 29
- Strauss M. A., Yahil A., Davis M., Huchra J. P., Fisher K., 1992b, ApJ, 397, 395
- Vogele M. S., Park C., Geller M. J., Huchra J. P., Gott J. R., 1994, ApJ, 420, 525
- Weinberg D. H., 1988, PASP, 100, 1373
- Weinberg D. H., Gott J. R., Melott A. L., 1987, ApJ, 321, 2

APPENDIX A: CLUSTERING AMPLITUDE DETERMINATION

In this appendix we outline our method to calculate the variance $\sigma^2(\lambda)$ of the smoothed PSCz density field. This measurement is used to estimate the shot-noise influence on the genus amplitude, and to obtain an estimate of the genus amplitude of the *Gaussianized* version of the PSCz density field. Our approach is that of Springel & White (1998), which is a slightly modified version of the method proposed by Saunders et al. (1991).

We start with an unsmoothed density field derived by binning the galaxies on a fine mesh. We assume that the probability of finding a galaxy in a voxel i of volume δV_i is given by $p \propto \rho_i S_i \delta V_i$, where ρ_i is the density, and S_i the value of the selection function in the cell. For convenience, we assume $\rho = 1$ in the following. Then the probability density function that describes the distribution of the counts in cell i is given by

$$p_i(m) = \sum_{N=0}^{\infty} \frac{\bar{m}_i^N}{N!} e^{-\bar{m}_i} \delta^{(D)}(m - N). \quad (A1)$$

Here $\bar{m}_i = \rho_i S_i$ denotes the expected number of galaxies in the cell. Moments of the counts in this cell can be derived from the moment-generating function

$$M_{m_i}(t) = \int p_i(m) e^{mt} dm = \exp[\bar{m}_i(e^t - 1)] \quad (A2)$$

by differentiating

$$\bar{m}_i^n = \left. \frac{d^n M_{m_i}}{dt^n} \right|_{t=0} \quad (A3)$$

at zero lag.

We now focus on the smoothed density field

$$d_i = \sum_j w_{ij} \frac{m_j}{S_j}, \quad (A4)$$

where w_{ij} is the value of the effective smoothing kernel (i.e., taking into account the renormalization due to the ratio method) between cells i and j .

Because the \bar{m}_i are independent variables with respect to the sampling process, the moment-generating function for d_i is simply

$$M_{d_i}(t) = \prod_j M_{m_j} \left(\frac{w_{ij}}{S_j} t \right) = \exp \left[\sum_j \bar{m}_j (e^{t w_{ij}/S_j} - 1) \right].$$

Hence, d_i is an unbiased estimate of the smoothed underlying field $\hat{\rho}_i = \sum_j w_{ij} \rho_j$. An unbiased estimate of the mean density d is therefore given by

$$d = \frac{\sum_i g_i d_i}{\sum_i g_i} \quad (A5)$$

for arbitrary weights g_i . We choose $g_i = [\text{var}(d_i)]^{-1}$, which provides a minimum-variance estimator for d . To compute

$$\text{var}(d) = \langle \bar{d}^2 \rangle - \langle \bar{d} \rangle^2 \quad (A6)$$

we first average over sampling realizations (overbar) and then with respect to the density field ρ_i (angular brackets).

We now want to determine the moments

$$R_n = \langle (\hat{\rho}_i - 1)^n \rangle, \quad (A7)$$

and, in particular, $R_2 = \sigma^2$. Introducing the moment-generating function

$$M_{d_i - \bar{d}_i}(t) = \exp \left[\sum_j \bar{m}_j (e^{t w_{ij}/S_j} - 1) - \bar{d}_i t \right], \quad (A8)$$

the central moments are given by

$$D_i^{(n)} = \overline{(d_i - \bar{d}_i)^n} = \left. \frac{d^n M_{d_i - \bar{d}_i}(t)}{dt^n} \right|_{t=0}. \quad (A9)$$

Note that due to the rescaling we have $\bar{d}_i = \hat{\rho}_i$. We can now relate the moments of the measured field with those of the underlying density field by using a binomial expansion

$$(d_i - 1)^n = \sum_{k=0}^n \binom{n}{k} (d_i - \bar{d}_i)^k (\hat{\rho}_i - 1)^{n-k}. \quad (A10)$$

In particular, for $n=2$ we find $\overline{(d_i - 1)^2} = D_i^{(2)} + (\hat{\rho}_i - 1)^2$.

Then an unbiased estimate for the variance of the smoothed underlying density field is given by

$$R_2 = \frac{\sum_i h_i [(d_i - 1)^2 - \langle D_i^{(2)} \rangle]}{\sum_i h_i} \quad (\text{A11})$$

for arbitrary weights h_i . Again, we want to estimate R_2 with minimum variance by choosing the weights as

$$h_i = \frac{1}{\text{var} [(d_i - 1)^2]}. \quad (\text{A12})$$

In order to compute these weights, we again make use of equation (A10):

$$\begin{aligned} (\overline{d_i - 1})^4 &= D_i^{(4)} + 4D_i^{(3)}(\hat{\rho}_i - 1)^2 \\ &\quad + 6D_i^{(2)}(\hat{\rho}_i - 1)^2 + (\hat{\rho}_i - 1)^4. \end{aligned} \quad (\text{A13})$$

Here the moments $D_i^{(n)}$ are given by equation (A9) as

$$\begin{aligned} D_i^{(2)} &= \sum_j \rho_j \frac{w_{ij}^2}{S_j}, & D_i^{(3)} &= \sum_j \rho_j \frac{w_{ij}^3}{S_j^2}, \\ D_i^{(4)} &= 3 \left(\sum_j \rho_j \frac{w_{ij}^2}{S_j} \right)^2 + \sum_j \rho_j \frac{w_{ij}^3}{S_j^2}. \end{aligned} \quad (\text{A14})$$

We now approximate the sums in these expressions by replacing the density ρ_j by the averaged (smoothed) density $\hat{\rho}_i$ in the particular region. We then obtain

$$\begin{aligned} D_i^{(2)} &= \hat{\rho}_i Y_i^{(2)}, & D_i^{(3)} &= \hat{\rho}_i Y_i^{(3)}, \\ D_i^{(4)} &= 3[\hat{\rho}_i Y_i^{(2)}]^2 + \hat{\rho}_i Y_i^{(4)}, \end{aligned} \quad (\text{A15})$$

where $Y_i^{(n)}$ is given by

$$Y_i^{(n)} = \sum_j \frac{w_{ij}^n}{S_j^{n-1}}. \quad (\text{A16})$$

Of course, this is only strictly correct for a homogeneous density field, but in the context of this analysis it can be expected to be a good approximation. Note that even if the weights h_i are obtained only approximately, the final estimate of R_2 remains unbiased.

We can now take the ensemble average of equation (A13) and compute the weights h_i . Here a well-known problem in counts-in-cells analysis arises: the determination of second and higher moments involves the whole hierarchy of moments. A minimum-variance estimate of R_2 requires the knowledge of R_3 and R_4 , which in turn depend on still higher moments. To close this hierarchy we assume that the density field is close to Gaussian, which implies $R_3 = 0$ and $R_4 = 3R_2^2$. With this assumption the weights become

$$h_i^{-1} = Y_i^{(4)} + 2[Y_i^{(2)}]^2 + 2R_2^2 + R_2 \{3[Y_i^{(2)}]^2 + 4Y_i^{(3)} + 4Y_i^{(2)}\}. \quad (\text{A17})$$

An estimate of R_2 can now be obtained by iteratively solving equation (A11).

Since the d_i are strongly correlated an estimate of the statistical error in the determination of R_2 by directly computing the variance of R_2 in the above formalism is unrealistic. Instead, we estimate the uncertainty in R_2 by using an ensemble of PSCz-like mock catalogues extracted from a N -body simulation.

APPENDIX B: GENUS AMPLITUDE AND SHOT NOISE

The genus amplitude of a Gaussian random field depends only on the behaviour of the two-point correlation function of the smoothed density field near the origin. We therefore estimate $\xi(\mathbf{r})$ under the influence of shot noise, allowing the noise level to vary across the survey volume.

We start by considering a discrete field $m(\mathbf{r}) = \sum_i \delta^{(D)}(\mathbf{r} - \mathbf{r}_i)$ of points arising in a Poisson sampling process of some underlying density field $\delta_i(\mathbf{r})$.

Allowing for a radial variation of the expected number density $S(\mathbf{r}) = \langle m(\mathbf{r}) \rangle$ of tracers, we define the measured density contrast as

$$\delta(\mathbf{r}) = \frac{m(\mathbf{r})}{S(\mathbf{r})} - 1. \quad (\text{B1})$$

This gives rise to a smoothed density field

$$\delta(\mathbf{r}) = \int W(\mathbf{r} - \mathbf{r}') \delta(\mathbf{r}') d\mathbf{r}', \quad (\text{B2})$$

with an autocorrelation function

$$\hat{\xi}(\mathbf{r}) = \langle \hat{\delta}(\mathbf{r}_0) \hat{\delta}(\mathbf{r}_0 + \mathbf{r}) \rangle. \quad (\text{B3})$$

Performing the ensemble average in equation (B3) with the Feldman, Kaiser & Peacock (1994) relation

$$\begin{aligned} \langle m(\mathbf{r}') m(\mathbf{r}'') \rangle &= S(\mathbf{r}') S(\mathbf{r}'') [1 + \xi_t(\mathbf{r}' - \mathbf{r}'')] \\ &\quad + S(\mathbf{r}') \delta^{(D)}(\mathbf{r}' - \mathbf{r}''), \end{aligned} \quad (\text{B4})$$

we find

$$\begin{aligned} \hat{\xi}(\mathbf{r}) &= \int W(\mathbf{r}') W(\mathbf{r} - \mathbf{r}' - \mathbf{r}') \xi_t(\mathbf{r}'') d\mathbf{r}' d\mathbf{r}'' \\ &\quad + \int W(\mathbf{r}_0 - \mathbf{r}') W(\mathbf{r}_0 + \mathbf{r} - \mathbf{r}') \frac{1}{S(\mathbf{r}')} d\mathbf{r}', \end{aligned} \quad (\text{B5})$$

where $\xi_t(\mathbf{r})$ is the autocorrelation function of the underlying field.

The second term in equation (B5) describes the shot-noise contribution, which varies across the survey volume. Hence the dependence on \mathbf{r}_0 remains even after the ensemble average. Assuming that $S(\mathbf{r})$ varies sufficiently slowly over a smoothing length, and choosing a Gaussian smoothing kernel for $W(\mathbf{r})$, we can substitute $S(\mathbf{r}')$ by $S(\mathbf{r}_0)$ in the integrand because $W(\mathbf{r}_0 - \mathbf{r}')$ is strongly peaked at \mathbf{r}_0 . We then obtain

$$\hat{\xi}(\mathbf{r}) = \hat{\xi}_t(\mathbf{r}) + \hat{\xi}_{\text{shot}}(\mathbf{r}), \quad (\text{B6})$$

where

$$\hat{\xi}_t(\mathbf{r}) = \frac{1}{(2\pi)^{3/2} \lambda^3} \int \xi_t(\mathbf{r}'') \exp \left[-\frac{(\mathbf{r} - \mathbf{r}'')^2}{2\lambda^2} \right] d\mathbf{r}'' \quad (\text{B7})$$

is the contribution of the underlying density field, and

$$\hat{\xi}_{\text{shot}}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2} \lambda^3 S(\mathbf{r}_0)} \exp \left(-\frac{\mathbf{r}^2}{2\lambda^2} \right) \quad (\text{B8})$$

stands for the shot-noise contribution.

Equation (B7) reads in Fourier space like

$$\hat{P}_t(\mathbf{k}) = P_t(\mathbf{k}) \exp\left(-\frac{\mathbf{k}^2 \lambda^2}{2}\right). \quad (\text{B9})$$

Adopting isotropy for the underlying field, we can also write

$$\hat{\xi}_t(r) = \int \hat{P}_t(k) \frac{\sin(kr)}{kr} d^3k. \quad (\text{B10})$$

We now assume that the density field can approximately be taken to be Gaussian. In this case the amplitude of the genus curve depends only on $\hat{\xi}(0)$ and $\hat{\xi}''(0)$ (Hamilton et al. 1986). The previous equations give the corresponding contributions from the underlying field as

$$\sigma_t^2 \equiv \hat{\xi}_t(0) = \int \hat{P}_t(k) d^3k \quad (\text{B11})$$

and

$$\hat{\xi}_t''(0) = \int \frac{k^2}{3} \hat{P}_t(k) d^3k = -\frac{\langle k^2 \rangle \sigma_t^2}{3}, \quad (\text{B12})$$

while the shot noise gives rise to

$$\sigma_{\text{shot}}^2 \equiv \hat{\xi}_{\text{shot}}(0) = \frac{1}{(2\pi)^{3/2} \lambda^3 S(r_0)} \quad (\text{B13})$$

and

$$\hat{\xi}_{\text{shot}}''(0) = -\frac{1}{(2\pi)^{3/2} \lambda^5 S(r_0)} = -\frac{\sigma_{\text{shot}}^2}{\lambda^2}. \quad (\text{B14})$$

We can now estimate the amplitude of the genus curve under the influence of shot noise. It is given by

$$\begin{aligned} N &= \frac{1}{(2\pi)^2} \left[\frac{-\hat{\xi}''(0)}{\hat{\xi}(0)} \right]^{3/2} \\ &= \frac{1}{(2\pi)^2} \left[\frac{\langle k^2 \rangle}{3} + \frac{\sigma_{\text{shot}}^2}{\sigma_t^2 + \sigma_{\text{shot}}^2} \left(\frac{1}{\lambda^2} - \frac{\langle k^2 \rangle}{3} \right) \right]^{3/2}. \end{aligned} \quad (\text{B15})$$